Response surface methodology: Asymptotic normality of the optimal solution

José A. Díaz-García *
Department of Statistics and Computation
25350 Buenavista, Saltillo, Coahuila, Mexico
E-mail: jadiaz@uaaan.mx

José E. Rodríguez
Department of Mathematics
University of Guanajuato
Jalisco s/n, Mineral de Valenciana
36240 Guanajuato, México
E-mail: elias.rodriguez@ugto.mx

Rogelio Ramos-Quiroga
Centro de Investigación en Matemáticas
Department of Probability and Statistics
Callejón de Jalisco s/n, Mineral de Valenciana
36240 Guanajuato, Guanajuato, México
E-mail: rramosq@cimat.mx

Abstract

Sensitivity analysis of the optimal solution in response surface methodology is studied and an explicit form of the effect of perturbation of the regression coefficients on the optimal solution is obtained. The characterisation of the critical point of the convex program corresponding to the optimum of a response surface model is also studied. The asymptotic normality of the optimal solution follows by standard methods.

1 Introduction

From a point of view of the mathematical programming, the sensitivity analysis studies the effect of small perturbations in the parameters on the optimal objective function value and on the critical point for mathematical programming problems. In general, these parameters shape the objective function and constraint the approach to the problem of mathematical programming. The sensitivity analysis of the mathematical programming has been studied by several authors, see Jagannathan (1977), Dupačová (1984) and Fiacco and Ghaemi (1982) among many other. As a direct consequence of the sensitivity analysis, the asymptotic

*Corresponding author
Key words. Multiple Response Surfaces (MRS), multiobjective optimisation, probabilistic optimisation, stochastic programming, matrix optimisation.
2000 Mathematical Subject Classification. 62K20; 90C15; 90C29

normality study of the critical point follows by standard methods of mathematical statistics (see similar results for the case of maximum likelihood estimates Aitchison and Silvey (1958)). This last consequence makes the sensitivity analysis very appealing for researching[1] from a statistical point of view.

In this work, we study the effect of perturbations of the regression parameters on the optimal solution of the response surface model and then the asymptotic normality of the critical point is obtained.

A very useful statistical tool in the study of designs, phenomena and experiments, is the response surfaces methodology. Which enables to find an analytical relationship between the response and controlled variables through a process of continuous improvement and optimisation.

Specifically, it is assumed that a researcher knows a system, which considers some observable response variables y which depends on some input variables, $x_1, \dots x_k$. Also it is assumed that the input variables $x_i^{'s}$ can be controlled by the researcher with a minimum error.

In general we have that

$$y(\mathbf{x}) = \eta(x_1, \dots x_k) \tag{1}$$

where the form of the function $\eta(\cdot)$, usually termed as the true response surface, is unknown and perhaps, very complicated; and $\mathbf{x} = (x_1, \dots x_k)'$. The success of the response surfaces methodology depends on the approximation of $\eta(\cdot)$ for a polynomial of low degree in some region.

In this work we assume that $\eta(\cdot)$ can be soundly approximated by a polynomial of second order, that is

$$y(\mathbf{x}) = \beta_0 + \sum_{i=1}^{n} \beta_i x_i + \sum_{i=1}^{n} \beta_{ii} x_i^2 + \sum_{i=1}^{n} \sum_{j>i}^{n} \beta_{ij} x_i x_j \tag{2}$$

where the unknown parameters $\beta_j^{'s}$ can be estimated via regression's techniques, as it is described in next section.

Next, we are interested in obtaining the levels of the input variables $x_i^{'s}$ such that the response variables y is minimum (optimal). This can be achieved if the following mathematical program is solved

$$\begin{array}{c} \min\limits_{\mathbf{x}} y(\mathbf{x}) \\ \text{subject to} \\ \mathbf{x} \in \mathfrak{X}, \end{array} \tag{3}$$

where $\mathfrak{X}$ is certain operating region for the input variables $x_i^{'s}$.

Now, two questions, intimately related, can be observed:

1. When the estimation of (2) is considered into (3) the critical point $\mathbf{x}^*$ obtained as solution shall be a function of the estimators $\widehat{\beta}_j^{'s}$ of the $\beta_j^{'s}$. Thus, given that $\widehat{\beta}_j^{'s}$ are random variables, then $\mathbf{x}^* \equiv \mathbf{x}^*(\widehat{\beta}_j^{'s})$ is a random vector too. The question is, if the distribution of $\widehat{\boldsymbol{\beta}}$ is known, then what is the distribution of $\mathbf{x}^*(\widehat{\beta}_j^{'s})$?

2. And, perhaps it is not sufficient to know only a point estimate of $\mathbf{x}^*(\widehat{\beta}_j^{'s})$, should be more convenient to know an interval estimate.

The distribution of the critical point in response surface methodology was studied by Díaz García and Ramos-Quiroga (2001, 2002), when $y(\mathbf{x})$ is defined as an hyperplane.

[1]In the context of the mathematical statistical the sensitivity analysis consist in to study the ways in which the estimators of certain model are affected by omission of a particular set of variables or by the inclusion or omission of a particular observation or set of observations, see Chatterjee and Hadi (1988).

In this work we give a response to these two questions. First, in Section 2 some notation is established. In Section 3 the response surface optimisation problem is proposed. First-order and second-order Kuhn-Tucker conditions are stated characterising the critical point in Section 4. Finally, the asymptotic normality of a critical point is established in Section 5.

2 Notation

A detailed discussion of response surface methodology may be found in Khuri and Cornell (1987) and Myers *et al.* (2009). For convenience, their principal properties and usual notation shall be restated here.

Let N be the number of experimental runs. The response variable is measured for each setting of a group of n coded variables (also are termed factors) $x_1, x_2, \ldots, x_n$. We assume that the response variable can be modeled by *second order polynomials regression model* in terms of $x_i^{'s}$. Hence, the response model can be written as

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{4}$$

where $\mathbf{y} \in \Re^N$ is the vector of observations on the response variable, $\mathbf{X} \in \Re^{N \times p}$ is an matrix of rank p termed design or regression matrix, $p = 1 + n + n(n+1)/2$, $\boldsymbol{\beta} \in \Re^p$ is a vector of unknown constant parameters, and $\boldsymbol{\varepsilon} \in \Re^N$ is a random error vector such that $\boldsymbol{\varepsilon} \sim \mathcal{N}_N(\mathbf{0}, \sigma^2\mathbf{I}_N)$, i.e. $\boldsymbol{\varepsilon}$ has an *N-dimensional normal distribution* with $\mathrm{E}(\boldsymbol{\varepsilon}) = \mathbf{0}$ and $\mathrm{Cov}(\boldsymbol{\varepsilon}) = \sigma^2\mathbf{I}_N$. Then we have:

- $\mathbf{x} = (x_1, x_2, \ldots, x_n)'$: Vector of controllable variables or factors. Formally, to each factor $A, B, \ldots$ is associated an $x_i^{'s}$ variable

- $\widehat{\boldsymbol{\beta}}$: The least squares estimator of

  $$\boldsymbol{\beta} = (\beta_0, \beta_1, \ldots, \beta_n, \beta_{11}, \ldots, \beta_{nn}, \beta_{12}, \ldots, \beta_{(n-1)n})'$$

  given by

  $$\widehat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = (\widehat{\beta}_0, \widehat{\beta}_1, \ldots, \widehat{\beta}_n, \widehat{\beta}_{11}, \ldots, \widehat{\beta}_{nn}, \widehat{\beta}_{12}, \ldots, \widehat{\beta}_{(n-1)n})'.$$

  Furthermore, under the assumption that $\boldsymbol{\varepsilon} \sim \mathcal{N}_N(\mathbf{0}, \sigma^2\mathbf{I}_N)$, then $\widehat{\boldsymbol{\beta}} \sim \mathcal{N}_p\left(\boldsymbol{\beta}, \sigma^2(\mathbf{X}'\mathbf{X})^{-1}\right)$.

- $\mathbf{z}(\mathbf{x}) = (1, x_1, x_2, \ldots, x_n, x_1^2, x_2^2, \ldots, x_n^2, x_1x_2, x_1x_3 \ldots, x_{n-1}x_n)'$.

- $\widehat{\boldsymbol{\beta}}_1 = (\widehat{\beta}_1, \ldots, \widehat{\beta}_n)'$ and

  $$\widehat{\mathbf{B}} = \frac{1}{2}\begin{pmatrix} 2\widehat{\beta}_{11} & \widehat{\beta}_{12} & \cdots & \widehat{\beta}_{1n} \\ \widehat{\beta}_{21} & 2\widehat{\beta}_{22} & \cdots & \widehat{\beta}_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \widehat{\beta}_{n1} & \widehat{\beta}_{n2} & \cdots & 2\widehat{\beta}_{nn} \end{pmatrix}$$

- $$\begin{aligned} \widehat{y}(\mathbf{x}) &= \mathbf{z}'(\mathbf{x})\widehat{\boldsymbol{\beta}} \\ &= \widehat{\beta}_0 + \sum_{i=1}^{n} \widehat{\beta}_i x_i + \sum_{i=1}^{n} \widehat{\beta}_{ii} x_i^2 + \sum_{i=1}^{n}\sum_{j>i}^{n} \widehat{\beta}_{ij} x_i x_j \\ &= \widehat{\beta}_0 + \widehat{\boldsymbol{\beta}}_1'\mathbf{x} + \mathbf{x}'\widehat{\mathbf{B}}\mathbf{x} : \end{aligned}$$

  Estimated response surface or predictor equation at the point $\mathbf{x}$.

- $\widehat{\sigma}^2 = \dfrac{\mathbf{y}'(\mathbf{I}_N - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\mathbf{y}}{N-p}$: Estimator of variance σ^2 such that $(N-p)\widehat{\sigma}^2/\sigma^2$ has a χ^2-distribution with $(N-p)$ freedom degrees; denoted this fact as $(N-p)\widehat{\sigma}^2/\sigma^2 \sim \chi^2_{N-p}$. Finally, note that

$$\widehat{\text{Cov}}(\widehat{\boldsymbol{\beta}}) = \widehat{\sigma}^2(\mathbf{X}'\mathbf{X})^{-1}. \tag{5}$$

3 Response surface optimisation

For convenience, those concepts and notations required here are listed below in terms of the estimated model of response surface. Definitions and details properties, etc., may be found in Khuri and Cornell (1987) and Myers *et al.* (2009).

Ideally, the goal would be to solve (3), however note that in general the parameters of form $\beta_j^{'s}$ are unknown. Then proceeding as it is indicated in the previous section, the $\beta_j^{'s}$ parameters are estimated. Then, the response surface optimisation problem in general is proposed as the following substitute mathematical program

$$\begin{array}{c} \min\limits_{\mathbf{x}} \widehat{y}(\mathbf{x}) \\ \text{subject to} \\ \mathbf{x} \in \mathfrak{X}, \end{array} \tag{6}$$

which is a nonlinear mathematical problem, see Khuri and Cornell (1987), and Rao (1979). Here $\mathfrak{X}$ denotes the experimental region, which is defined, in general, as a hypercube

$$\mathfrak{X} = \{\mathbf{x} | l_i < x_i < u_i, \quad i = 1, 2, \dots, n\},$$

where $\mathbf{l} = (l_1, l_2, \dots, l_n)'$ defines the vector of lower bounds of factors, and $\mathbf{u} = (u_1, u_2, \dots, u_n)'$ defines the vector of upper bounds of factors. Alternatively, the experimental region is defined as an hypersphere

$$\mathfrak{X} = \{\mathbf{x} | \mathbf{x}'\mathbf{x} \leq c^2, c > 0\}, \tag{7}$$

where note that $\mathbf{x}'\mathbf{x} = ||\mathbf{x}||^2$ and in general c is specified by the experimental design model used. In this paper it is considered $\mathfrak{X}$ defined by (7).

4 Characterisation of the critical point

Let $\mathbf{x}^*(\widehat{\boldsymbol{\beta}}) \in \Re^n$ be the unique optimal solution of program (6) with the corresponding Lagrange multiplier $\lambda^*(\widehat{\boldsymbol{\beta}}) \in \Re$. The Lagrangian is defined by

$$L(\mathbf{x}, \lambda; \widehat{\boldsymbol{\beta}}) = \widehat{y}(\mathbf{x}) + \lambda(||\mathbf{x}||^2 - c^2). \tag{8}$$

Similarly, $\mathbf{x}^*(\boldsymbol{\beta}) \in \Re^n$ denotes the unique optimal solution of program (3) with the corresponding Lagrange multiplier $\lambda^*(\boldsymbol{\beta}) \in \Re$.

Now we establish the local Kuhn-Tucker conditions that guarantee that the Kuhn-Tucker point $\mathbf{r}^*(\widehat{\boldsymbol{\beta}}) = \left[\mathbf{x}^*(\widehat{\boldsymbol{\beta}}), \lambda^*(\widehat{\boldsymbol{\beta}})\right]' \in \Re^{n+1}$ is a unique global minimum of convex program (6). First recall that for $f : \Re^n \to \Re$, then $\dfrac{\partial f}{\partial \mathbf{x}} \equiv \nabla_{\mathbf{x}}$ denotes the gradient of function f.

Theorem 4.1. *The necessary and sufficient conditions that a point* $\mathbf{x}^*(\widehat{\boldsymbol{\beta}}) \in \Re^n$ *for arbitrary fixed* $\widehat{\boldsymbol{\beta}} \in \Re^p$*, be a unique global minimum of the convex program (6) is that,* $\mathbf{x}^*(\widehat{\boldsymbol{\beta}})$ *and the*

corresponding Lagrange multiplier $\lambda^*(\widehat{\boldsymbol{\beta}}) \in \Re$*, fulfill the Kuhn-Tucker first order conditions*

$$
\begin{aligned}
\nabla_{\mathbf{x}} L(\mathbf{x}, \lambda; \widehat{\boldsymbol{\beta}}) = \left\{ \begin{array}{l} \mathbf{M}(\mathbf{x})\widehat{\boldsymbol{\beta}} + 2\lambda(\widehat{\boldsymbol{\beta}})\mathbf{x} \\ \quad or \\ \widehat{\boldsymbol{\beta}}_1 + 2(\widehat{\mathbf{B}} + \lambda(\widehat{\boldsymbol{\beta}})\mathbf{I}_n)\mathbf{x} \end{array} \right\} &= \mathbf{0} && (9)\\
\nabla_{\lambda} L(\mathbf{x}, \lambda; \widehat{\boldsymbol{\beta}}) = ||\mathbf{x}||^2 - c^2 &\leq 0 && (10)\\
\lambda(\widehat{\boldsymbol{\beta}})(||\mathbf{x}||^2 - c^2) &= 0 && (11)\\
\lambda(\widehat{\boldsymbol{\beta}}) &\geq 0 && (12)
\end{aligned}
$$

where

$$
\begin{aligned}
\mathbf{M}(\mathbf{x}) &= \nabla_{\mathbf{x}} \mathbf{z}'(\mathbf{x}) = \frac{\partial \mathbf{z}'(\mathbf{x})}{\partial \mathbf{x}} \\
&= (\mathbf{0} \vdots \mathbf{I}_n \vdots 2\operatorname{diag}(\mathbf{x}) \vdots \mathbf{C}_1 \vdots \cdots \vdots \mathbf{C}_{n-1}) \in \Re^{n \times p},
\end{aligned}
$$

with

$$
\mathbf{C}_i = \begin{pmatrix} \mathbf{0}'_1 \\ \vdots \\ \mathbf{0}'_{i-1} \\ \mathbf{x}'\mathbf{A}_i \\ x_i \mathbf{I}_{n-i} \end{pmatrix}, i = 1, \dots, n-1, \quad \mathbf{0}_j \in \Re^{n-i}, j = 1, \dots, i-1,
$$

observing that when $i = 1$ *(i.e.* $j = 0$*) means that this row not appear in* $\mathbf{C}_1$*; and*

$$
\mathbf{A}_i = \begin{pmatrix} \mathbf{0}'_1 \\ \vdots \\ \mathbf{0}'_i \\ \mathbf{I}_{n-i} \end{pmatrix}, \quad \mathbf{0}'_k \in \Re^{n-i}, k = 1, \dots, i.
$$

In addition, assume that strict complementarity slackness holds at $\mathbf{x}^*(\boldsymbol{\beta})$ with respect to $\lambda^*(\boldsymbol{\beta})$, that is

$$
\lambda^*(\boldsymbol{\beta}) > 0 \Leftrightarrow ||\mathbf{x}||^2 - c^2 = 0. \tag{13}
$$

Analogously, the Khun-Tucker condition (9) to (12) for $\widehat{\boldsymbol{\beta}} = \boldsymbol{\beta}$ are stated next.

Corollary 4.1. *The necessary and sufficient conditions that a point* $\mathbf{x}^*(\boldsymbol{\beta}) \in \Re^n$ *for arbitrary fixed* $\boldsymbol{\beta} \in \Re^p$*, be a unique global minimum of the convex program (3) is that,* $\mathbf{x}^*(\boldsymbol{\beta})$ *and the corresponding Lagrange multiplier* $\lambda^*(\boldsymbol{\beta}) \in \Re$*, fulfill the Kuhn-Tucker first order conditions*

$$
\begin{aligned}
\nabla_{\mathbf{x}} L(\mathbf{x}, \lambda; \widehat{\boldsymbol{\beta}}) = \left\{ \begin{array}{l} \mathbf{M}(\mathbf{x})\boldsymbol{\beta} + 2\lambda(\boldsymbol{\beta})\mathbf{x} \\ \quad or \\ \boldsymbol{\beta}_1 + 2(\mathbf{B}\mathbf{x} + \lambda(\boldsymbol{\beta})\mathbf{I}_n)\mathbf{x} \end{array} \right\} &= \mathbf{0} && (14)\\
\nabla_{\lambda} L(\mathbf{x}, \lambda; \boldsymbol{\beta}) = ||\mathbf{x}||^2 - c^2 &\leq 0 && (15)\\
\lambda(\boldsymbol{\beta})(||\mathbf{x}||^2 - c^2) &= 0 && (16)\\
\lambda(\boldsymbol{\beta}) &\geq 0 && (17)
\end{aligned}
$$

and $\lambda(\boldsymbol{\beta}) = 0$ *when* $||\mathbf{x}||^2 - c^2 < 0$ *at* $[\mathbf{x}^*(\boldsymbol{\beta}), \lambda^*(\boldsymbol{\beta})]'$.

Observe that, due to the strict convexity of the constraint and objective function, the second-order sufficient condition is evidently fulfilled for the convex program (6).

Next is established the existence of a once continuously differentiable solution to program (6), see Fiacco and Ghaemi (1982).

Theorem 4.2. *Assume that (13) hold true and the second-order sufficient condition is for the convex program (6). Then*

1. $\mathbf{x}^*(\boldsymbol{\beta})$ *is a unique global minimum of program (3) and* $\lambda^*(\boldsymbol{\beta})$ *is unique too.*

2. *For* $\widehat{\boldsymbol{\beta}} \in V_\varepsilon(\boldsymbol{\beta})$ *(is an* $\varepsilon-$*neighborhood or open ball), there exist a unique once continuously differentiable vector function*

$$\mathbf{r}^*(\widehat{\boldsymbol{\beta}}) = \begin{bmatrix} \mathbf{x}^*(\widehat{\boldsymbol{\beta}}) \\ \lambda^*(\widehat{\boldsymbol{\beta}}) \end{bmatrix} \in \Re^{n+1}$$

satisfying the second order sufficient conditions of problem (3)such that $\mathbf{r}^*(\boldsymbol{\beta}) = [\mathbf{x}^*(\boldsymbol{\beta}), \lambda^*(\boldsymbol{\beta})]'$ *and hence,* $\mathbf{x}^*(\widehat{\boldsymbol{\beta}})$ *is a unique global minimum of problem (6) with associated unique Lagrange multiplier* $\lambda^*(\widehat{\boldsymbol{\beta}})$.

3. *For* $\widehat{\boldsymbol{\beta}} \in V_\varepsilon(\boldsymbol{\beta})$*, the status of the constraint is unchanged and* $\lambda^*(\widehat{\boldsymbol{\beta}}) > 0 \Leftrightarrow ||\mathbf{x}||^2 - c^2 = 0$ *is hold.*

5 Asymptotic normality of the critical point

This section considers the statistical and mathematical programming aspects of the sensitivity analysis of the optimum of a estimated response surface model.

Theorem 5.1. *Assume:*

1. *For any* $\widehat{\boldsymbol{\beta}} \in V_\varepsilon(\boldsymbol{\beta})$*, the second-order sufficient condition is fulfilled for the convex program (6) such that the second order derivatives*

$$\frac{\partial^2 L(\mathbf{x}, \lambda; \widehat{\boldsymbol{\beta}})}{\partial \mathbf{x} \partial \mathbf{x}'}, \ \frac{\partial^2 L(\mathbf{x}, \lambda; \widehat{\boldsymbol{\beta}})}{\partial \mathbf{x} \partial \widehat{\boldsymbol{\beta}}'}, \ \frac{\partial^2 L(\mathbf{x}, \lambda; \widehat{\boldsymbol{\beta}})}{\partial \mathbf{x} \partial \lambda}, \ \frac{\partial^2 L(\mathbf{x}, \lambda; \widehat{\boldsymbol{\beta}})}{\partial \lambda \partial \mathbf{x}'}, \ \frac{\partial^2 L(\mathbf{x}, \lambda; \widehat{\boldsymbol{\beta}})}{\partial \lambda \partial \widehat{\boldsymbol{\beta}}}$$

exist and are continuous in $\left[\mathbf{x}^*(\widehat{\boldsymbol{\beta}}), \lambda^*(\widehat{\boldsymbol{\beta}})\right]' \in V_\varepsilon([\mathbf{x}^*(\boldsymbol{\beta}), \lambda^*(\boldsymbol{\beta})]')$ *and*

$$\frac{\partial^2 L(\mathbf{x}, \lambda; \widehat{\boldsymbol{\beta}})}{\partial \mathbf{x} \partial \mathbf{x}'},$$

is positive definite.

2. $\widehat{\boldsymbol{\beta}}_\nu$ *the estimator of the true parameter vector* $\boldsymbol{\beta}_\nu$ *that is based on a sample of size* N_ν *is such that*

$$\sqrt{N_\nu}(\widehat{\boldsymbol{\beta}}_\nu - \boldsymbol{\beta}_\nu) \sim \mathcal{N}_p(\mathbf{0}_p, \boldsymbol{\Sigma}), \quad \frac{1}{N_\nu}\boldsymbol{\Sigma} = \sigma^2(\mathbf{X}'\mathbf{X})^{-1}.$$

3. *(13) is fulfilled for* $\widehat{\boldsymbol{\beta}} = \boldsymbol{\beta}$*. Then asymptotically*

$$\sqrt{N_\nu}\left[\mathbf{x}^*(\widehat{\boldsymbol{\beta}}) - \mathbf{x}^*(\boldsymbol{\beta})\right] \xrightarrow{d} \mathcal{N}_n(\mathbf{0}_n, \boldsymbol{\Xi})$$

where the $n \times n$ *variance-covariance matrix*

$$\boldsymbol{\Xi} = \left(\frac{\partial \mathbf{x}^*(\widehat{\boldsymbol{\beta}})}{\partial \widehat{\boldsymbol{\beta}}}\right) \widehat{\boldsymbol{\Sigma}} \left(\frac{\partial \mathbf{x}^*(\widehat{\boldsymbol{\beta}})}{\partial \widehat{\boldsymbol{\beta}}}\right)',$$

such that all elements of $\left(\partial \mathbf{x}^*(\widehat{\boldsymbol{\beta}})/\partial \widehat{\boldsymbol{\beta}}\right)$ *are continuous on any* $\widehat{\boldsymbol{\beta}} \in V_\varepsilon(\boldsymbol{\beta})$*; furthermore*

$$\left(\frac{\partial \mathbf{x}^*(\widehat{\boldsymbol{\beta}})}{\partial \widehat{\boldsymbol{\beta}}}\right) = \mathbf{G}^{-1}\left(\frac{\mathbf{x}^*(\widehat{\boldsymbol{\beta}})\mathbf{x}^*(\widehat{\boldsymbol{\beta}})'\mathbf{G}^{-1}}{\mathbf{x}^*(\widehat{\boldsymbol{\beta}})'\mathbf{G}^{-1}\mathbf{x}^*(\widehat{\boldsymbol{\beta}})} - \mathbf{I}_n\right)\mathbf{M}\left(\mathbf{x}^*(\widehat{\boldsymbol{\beta}})\right),$$

where

$$\mathbf{G} = \frac{\partial^2 L(\mathbf{x}, \lambda; \widehat{\boldsymbol{\beta}})}{\partial \mathbf{x} \partial \mathbf{x}'} = 2\left(\widehat{\mathbf{B}} - \lambda^*(\widehat{\boldsymbol{\beta}})\mathbf{I}_n\right).$$

Proof. According to Theorem 4.1 and Corollary 4.1, the Kuhn-Tucker conditions (9)–(12) at $\left[\mathbf{x}^*(\widehat{\boldsymbol{\beta}}), \lambda^*(\widehat{\boldsymbol{\beta}})\right]'$ and to at (14)–(17) $[\mathbf{x}^*(\boldsymbol{\beta}), \lambda^*(\boldsymbol{\beta})]'$ are fulfilled for mathematical programs (3) and (6), respectively. From conditions (14)–(17) of Corollary 4.1, the following system equation

$$\nabla_{\mathbf{x}} L(\mathbf{x}, \lambda; \widehat{\boldsymbol{\beta}}) = \left\{\begin{array}{c} \mathbf{M}(\mathbf{x})\boldsymbol{\beta} + 2\lambda(\boldsymbol{\beta})\mathbf{x} \\ \text{or} \\ \boldsymbol{\beta}_1 + 2(\mathbf{B}\mathbf{x} + \lambda(\boldsymbol{\beta})\mathbf{I}_n)\mathbf{x} \end{array}\right\} = \mathbf{0} \tag{18}$$

$$\nabla_\lambda L(\mathbf{x}, \lambda; \boldsymbol{\beta}) = ||\mathbf{x}||^2 - c^2 = 0 \tag{19}$$

has a solution $\mathbf{x}^*(\boldsymbol{\beta}), \lambda^*(\boldsymbol{\beta}) > 0, \boldsymbol{\beta}$.

The nonsingular Jacobian matrix of the continuously differentiable functions (18) and (19) with respect to $\mathbf{x}$ and λ at $\left[\mathbf{x}^*(\widehat{\boldsymbol{\beta}}), \lambda^*(\widehat{\boldsymbol{\beta}})\right]'$ is

$$\left(\begin{array}{cc} \dfrac{\partial^2 L(\mathbf{x}, \lambda; \widehat{\boldsymbol{\beta}})}{\partial \mathbf{x}\partial \mathbf{x}'} & \dfrac{\partial^2 L(\mathbf{x}, \lambda; \widehat{\boldsymbol{\beta}})}{\partial \lambda \partial \mathbf{x}} \\ \dfrac{\partial^2 L(\mathbf{x}, \lambda; \widehat{\boldsymbol{\beta}})}{\partial \mathbf{x}'\partial \lambda} & 0 \end{array}\right) = \left(\begin{array}{cc} 2(\widehat{\mathbf{B}} + \lambda \mathbf{x}) & 2\mathbf{x} \\ 2\mathbf{x}' & 0 \end{array}\right).$$

According to the implicit functions theorem, there is a neighborhood $V_\varepsilon(\boldsymbol{\beta})$ such that for arbitrary $\widehat{\boldsymbol{\beta}} \in V_\varepsilon(\boldsymbol{\beta})$, the system (18) and (19) has a unique solution $\mathbf{x}^*(\widehat{\boldsymbol{\beta}})$, $\lambda^*(\widehat{\boldsymbol{\beta}})$, $\widehat{\boldsymbol{\beta}}$ and by Theorem 4.2, the components of $\mathbf{x}^*(\widehat{\boldsymbol{\beta}})$, $\lambda^*(\widehat{\boldsymbol{\beta}})$ are continuously differentiable function of $\widehat{\boldsymbol{\beta}}$, see Bigelow and Shapiro (1974). Their derivatives are given by

$$\begin{aligned} \left(\begin{array}{c} \dfrac{\partial \mathbf{x}^*(\widehat{\boldsymbol{\beta}})}{\partial \widehat{\boldsymbol{\beta}}} \\ \dfrac{\partial \lambda^*(\widehat{\boldsymbol{\beta}})}{\partial \widehat{\boldsymbol{\beta}}} \end{array}\right) &= -\left(\begin{array}{cc} \dfrac{\partial^2 L(\mathbf{x}, \lambda; \widehat{\boldsymbol{\beta}})}{\partial \mathbf{x}\partial \mathbf{x}'} & \dfrac{\partial^2 L(\mathbf{x}, \lambda; \widehat{\boldsymbol{\beta}})}{\partial \lambda \partial \mathbf{x}} \\ \dfrac{\partial^2 L(\mathbf{x}, \lambda; \widehat{\boldsymbol{\beta}})}{\partial \mathbf{x}'\partial \lambda} & 0 \end{array}\right)^{-1} \left(\begin{array}{c} \dfrac{\partial^2 L(\mathbf{x}, \lambda; \widehat{\boldsymbol{\beta}})}{\partial \mathbf{x}\partial \widehat{\boldsymbol{\beta}}'} \\ 0 \end{array}\right) \\ &= -\left(\begin{array}{cc} 2(\widehat{\mathbf{B}} + \lambda \mathbf{x}) & 2\mathbf{x} \\ 2\mathbf{x}' & 0 \end{array}\right)^{-1} \left(\begin{array}{c} \mathbf{M}(\mathbf{x}) \\ 0 \end{array}\right). \end{aligned} \tag{20}$$

The explicit form of $(\partial \mathbf{x}^*(\widehat{\boldsymbol{\beta}})/\partial \widehat{\boldsymbol{\beta}})$ follow from (20) and by formula

$$\left(\begin{array}{cc} \mathbf{P} & \mathbf{Q} \\ \mathbf{Q}' & \mathbf{0} \end{array}\right)^{-1} = \left(\begin{array}{cc} [\mathbf{I} - \mathbf{P}^{-1}\mathbf{Q}(\mathbf{Q}'\mathbf{P}^{-1}\mathbf{Q})^{-1}\mathbf{Q}']\mathbf{P}^{-1} & \mathbf{P}^{-1}\mathbf{Q}(\mathbf{Q}'\mathbf{P}^{-1}\mathbf{Q})^{-1} \\ (\mathbf{Q}'\mathbf{P}^{-1}\mathbf{Q})^{-1}\mathbf{Q}'\mathbf{P}^{-1} & -(\mathbf{Q}'\mathbf{P}^{-1}\mathbf{Q})^{-1} \end{array}\right)$$

where $\mathbf{P}$ and $\mathbf{S}$ are symmetric and $\mathbf{P}$ is nonsingular.

Then from assumption 2 and Rao (1973, (iii), p. 388) and Bishop *et al.* (1991, Theorem 14.6-2, p. 493) (see also Cramér (1946, p. 353))

$$\sqrt{N_\nu}\left[\mathbf{x}^*(\widehat{\boldsymbol{\beta}}) - \mathbf{x}^*(\boldsymbol{\beta})\right] \xrightarrow{d} \mathcal{N}_n\left(\mathbf{0}_n, \left(\frac{\partial \mathbf{x}^*(\boldsymbol{\beta})}{\partial \widehat{\boldsymbol{\beta}}}\right)\mathbf{\Sigma}\left(\frac{\partial \mathbf{x}^*(\boldsymbol{\beta})}{\partial \widehat{\boldsymbol{\beta}}}\right)'\right). \tag{21}$$

Finally note that all elements of $(\partial \mathbf{x}^*/\partial \widehat{\boldsymbol{\beta}})$ are continuous on $V_\varepsilon(\boldsymbol{\beta})$, so that the asymptotical distribution (21) can be substituted by

$$\sqrt{N_\nu}\left[\mathbf{x}^*(\widehat{\boldsymbol{\beta}}) - \mathbf{x}^*(\boldsymbol{\beta})\right] \xrightarrow{d} \mathcal{N}_n\left(\mathbf{0}_n, \left(\frac{\partial \mathbf{x}^*(\widehat{\boldsymbol{\beta}})}{\partial \widehat{\boldsymbol{\beta}}}\right) \widehat{\boldsymbol{\Sigma}} \left(\frac{\partial \mathbf{x}^*(\widehat{\boldsymbol{\beta}})}{\partial \widehat{\boldsymbol{\beta}}}\right)'\right),$$

see Rao (1973, (iv), pp.388–389). $\square$ $\square$

Now, consider that $\lambda(\boldsymbol{\beta}) = 0$ that is, $||\mathbf{x}||^2 - c^2 < 0$ at $\mathbf{x}^*(\widehat{\boldsymbol{\beta}})$. Thus in this case we have an explicit expression for $\mathbf{x}^*(\widehat{\boldsymbol{\beta}})$, furthermore

$$\mathbf{x}^*(\widehat{\boldsymbol{\beta}}) = -\frac{1}{2}\widehat{\mathbf{B}}^{-1}\widehat{\mathbf{b}}_1.$$

Hence:

Corollary 5.1. *Assume:*

1. *For any $\widehat{\boldsymbol{\beta}} \in V_\varepsilon(\boldsymbol{\beta})$, the second-order sufficient condition is fulfilled for the convex program (6) such that the second order derivatives*

$$\frac{\partial^2 \widehat{y}(\mathbf{x})}{\partial \mathbf{x} \partial \mathbf{x}'}, \ \frac{\partial^2 \widehat{y}(\mathbf{x})}{\partial \mathbf{x} \partial \widehat{\boldsymbol{\beta}}'},$$

   exist and are continuous in $\mathbf{x}^*(\widehat{\boldsymbol{\beta}}) \in V_\varepsilon(\mathbf{x}^*(\boldsymbol{\beta}))$ and

$$\frac{\partial^2 \widehat{y}(\mathbf{x})}{\partial \mathbf{x} \partial \mathbf{x}'},$$

   is positive definite.

2. *$\widehat{\boldsymbol{\beta}}_\nu$ the estimator of the true parameter vector $\boldsymbol{\beta}_\nu$ that is based on a sample of size N_ν is such that*

$$\sqrt{N_\nu}(\widehat{\boldsymbol{\beta}}_\nu - \boldsymbol{\beta}_\nu) \sim \mathcal{N}_p(\mathbf{0}_p, \boldsymbol{\Sigma}), \quad \frac{1}{N_\nu}\boldsymbol{\Sigma} = \sigma^2(\mathbf{X}'\mathbf{X})^{-1}.$$

3. *Then asymptotically*

$$\sqrt{N_\nu}\left[\mathbf{x}^*(\widehat{\boldsymbol{\beta}}) - \mathbf{x}^*(\boldsymbol{\beta})\right] \xrightarrow{d} \mathcal{N}_n(\mathbf{0}_n, \boldsymbol{\Xi}_1)$$

   where the $n \times n$ variance-covariance matrix

$$\boldsymbol{\Xi}_1 = \left(\frac{\partial \mathbf{x}^*(\widehat{\boldsymbol{\beta}})}{\partial \widehat{\boldsymbol{\beta}}}\right) \widehat{\boldsymbol{\Sigma}} \left(\frac{\partial \mathbf{x}^*(\widehat{\boldsymbol{\beta}})}{\partial \widehat{\boldsymbol{\beta}}}\right)',$$

   such that all elements of $\left(\partial \mathbf{x}^*(\widehat{\boldsymbol{\beta}})/\partial \widehat{\boldsymbol{\beta}}\right) \in \Re^{n \times p}$ are continuous on any $\widehat{\boldsymbol{\beta}} \in V_\varepsilon(\boldsymbol{\beta})$; furthermore

$$\begin{aligned}\left(\frac{\partial \mathbf{x}^*(\widehat{\boldsymbol{\beta}})}{\partial \widehat{\boldsymbol{\beta}}}\right) &= \left(\frac{\partial^2 \widehat{y}(\mathbf{x})}{\partial \mathbf{x} \partial \mathbf{x}'}\right)^{-1} \frac{\partial^2 \widehat{y}(\mathbf{x})}{\partial \mathbf{x} \partial \widehat{\boldsymbol{\beta}}'}, \\ &= \frac{1}{2}\widehat{\mathbf{B}}^{-1}\mathbf{M}\left(\mathbf{x}^*(\widehat{\boldsymbol{\beta}})\right).\end{aligned}$$

Proof. This is a verbatim copy of the proof of Theorem 4.2, noting that the conditions (14)–(17) are simply reduced to

$$\nabla_{\mathbf{x}} y(\mathbf{x}) = \left\{ \begin{array}{c} \mathbf{M}(\mathbf{x})\boldsymbol{\beta} \\ \text{or} \\ \boldsymbol{\beta}_1 + 2\mathbf{B}\mathbf{x} \end{array} \right\} = \mathbf{0},$$

which has the solution

$$\mathbf{x}^*(\boldsymbol{\beta}) = -\frac{1}{2}\mathbf{B}^{-1}\mathbf{b}_1. \qquad \square$$

$\square$

Conclusions

As consequence of Theorem 4.2 and Corollary 5.1 now is feasible to establish confidence intervals and hypothesis tests on the critical point, see Bishop *et al.* (1991, Section 14.6.4, pp. 498–500); then it is possible to establish operating conditions in intervals instead of isolated points.

Unfortunately in many applications, especially in industry, the number of observations is relatively small and perhaps the results obtained in this work should be applied with caution. However, the results of this paper can be taken as a good first approximation to the exact problem.

Acknowledgements

The first author was partially supported IDI-Spain, Grants No. FQM2006-2271 and MTM2008-05785. This paper was written during J. A. Díaz-García's stay as a professor at the Department of Mathematics of the University of Guanajuato.

References

J. Aitchison and S.D. Silvey, Maximum likelihoocl estimation of parameters subject to restraints, Ann. Mathe. Statist., 29 (1958), 813–828.

J. H. Gigelow, and N. Z. Shapiro, Implicit function theorem for mathematical progrtamming and for systems of iniqualities, Math. Programm. 6(2)(1974), 141–156.

Y. M. M. Bishop, S. E. Finberg, and P. W. Holland, Discrete Multivariate Analysis: Theory and Practice, The MIT press, Cambridge, 1991.

S. Chatterjee, and A. S. Hadi, Sensitivity Analysis in Linear Regression, John Wiley: New York, 1988.

H. Cramér, Mathematical Methods of Statistics, Princeton University Press, Princeton, 1946.

J. A. Díaz García, and R. Ramos-Quiroga, An approach to optimization in response surfaces, Comm. in Stat., Part A- Theory and Methods, 30(2001) 827–835.

J. A. Díaz García, and R. Ramos-Quiroga, Erratum. An approach to optimization in response surfaces, Comm. in Stat., Part A- Theory and Methods, 31(2002) 161.

J. Dupačová, Stability in stochastic programming with recourse-estimated parameters, Math. Prog., 28(1984) 72–83.

A. V. Fiacco and A. Ghaemi, Sensitivity analysis of a nonlinear structural design problem, Comp. and Ops. Res., 9(1)(1982) 29–55.

R. Jagannathan, Minimax procedure for a class of linear programs under uncertainty, Op. Res. 25(1977) 173–177.

A. I. Khuri, and J. A. Cornell, Response Surfaces: Designs and Analyses, Marcel Dekker, Inc., NewYork, 1987.

R. H. Myers, D. C. Montgomery, and C. M. Anderson-Cook, Response surface methodology: process and product optimization using designed experiments; third edition, Wiley, New York, 2009.

C. R. Rao, Linear Statistical Inference and its Applications (2nd ed.), John Wiley & Sons, New York, 1973.

S. S. Rao, Optimization Theory and Applications, Wiley Eastern Limited, New Delhi, 1979.